Math 310 **Fall 2023**
Summary of Course Content **Drew Armstrong**

Basic Vector Rules. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $r, s \in \mathbb{R}$.

$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

$\mathbf{u} + \mathbf{0} = \mathbf{u}$

$r(s\mathbf{u}) = (rs)\mathbf{u}$

$(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$

$r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$

$1\mathbf{u} = \mathbf{u}$

$0\mathbf{u} = \mathbf{0}$

Dot Product Rules. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

$\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}$

$\mathbf{u} \bullet (\mathbf{v} + \mathbf{w}) = \mathbf{u} \bullet \mathbf{v} + \mathbf{u} \bullet \mathbf{w}$

$c(\mathbf{u} \bullet \mathbf{v}) = (c\mathbf{u}) \bullet \mathbf{v} = \mathbf{u} \bullet (c\mathbf{v})$

$\mathbf{u} \bullet \mathbf{u} = \|\mathbf{u}\|^2$

$\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$

$\mathbf{u} \bullet \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$ [angle measured tail to tail]

$\mathbf{u} \bullet \mathbf{v} = 0$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are perpendicular

Cross Product Rules. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ and $c \in \mathbb{R}$.

$\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$

$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$

$c(\mathbf{u} \times \mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (c\mathbf{v})$

$\mathbf{u} \times \mathbf{0} = \mathbf{0}$

$\mathbf{u} \times \mathbf{u} = \mathbf{0}$

$\mathbf{u} \bullet (\mathbf{u} \times \mathbf{v}) = 0$

$\mathbf{v} \bullet (\mathbf{u} \times \mathbf{v}) = 0$

$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta$ [angle measured tail to tail, $0 \leq \theta < \pi$]

$\|\mathbf{u} \times \mathbf{v}\| =$ area of the parallelogram generated by $\mathbf{u}$ and $\mathbf{v}$

[There are few more fancy rules that you don't need to memorize.]

Lines in 2D. The line that passes through point (x_0, y_0) and is parallel to vector $\langle u, v\rangle$ can be parametrized as

$$\begin{aligned} \mathbf{r}(t) &= (x_0, y_0) + t\langle u, v\rangle \\ &= (x_0 + tu, y_0 + tv). \end{aligned}$$

The line that passes through the point (x_0, y_0) and is perpendicular to vector $\langle a, b\rangle$ has equation

$$\begin{gathered} \langle a, b\rangle \bullet \langle x - x_0, y - y_0\rangle = 0 \\ a(x - x_0) + b(y - y_0) = 0. \end{gathered}$$

Planes in 3D. The plane that passes through point (x_0, y_0, z_0) and is perpendicular to the vector $\langle a, b, c\rangle$ has equation

$$\begin{gathered} \langle a, b, c\rangle \bullet \langle x - x_0, y - y_0, z - z_0\rangle = 0 \\ a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \end{gathered}$$

Given three points in a plane one can use the cross product to find a normal vector.

Lines in 3D. The line that passes through point (x_0, y_0, z_0) and is parallel to the vector $\langle u, v, w\rangle$ can be parametrized as

$$\begin{aligned} \mathbf{r}(t) &= (x_0, y_0, z_0) + t\langle u, v, w\rangle \\ &= (x_0 + tu, y_0 + tv, z_0 + tw). \end{aligned}$$

This line cannot be described by just one equation. But it can be viewed as the solution of two simultaneous equations. (Geometrically, a line is an intersection of two planes.)

Derivative Product Rules. Let $\mathbf{u}, \mathbf{v} : \mathbb{R} \to \mathbb{R}^n$, $f : \mathbb{R} \to \mathbb{R}$, and $c \in \mathbb{R}$.

$[c\mathbf{u}(t)]' = c\mathbf{u}'(t)$

$[f(t)\mathbf{u}(t)]' = f'(t)\mathbf{u}(t) + f(t)\mathbf{u}'(t)$

$[\mathbf{u}(t) \bullet \mathbf{v}(t)]' = \mathbf{u}'(t) \bullet \mathbf{v}(t) + \mathbf{u}(t) \bullet \mathbf{v}'(t)$

$[\mathbf{u}(t) \times \mathbf{v}(t)]' = \mathbf{u}'(t) \times \mathbf{v}(t) + \mathbf{u}'(t) \times \mathbf{v}(t)$ [here we need $n = 3$]

Motion in Space. A function $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ can be viewed as the position of a particle at time t. The derivatives $\mathbf{r}'(t)$ and $\mathbf{r}''(t)$ are the velocity and acceleration vectors. The distance traveled between times t_1 and t_2 is the arc length

$$\begin{aligned} \text{distance} &= \int \text{speed} \\ \text{arc length} &= \int_{t_1}^{t_2} \|\mathbf{r}'(t)\|\, dt. \end{aligned}$$

If the particle $\mathbf{r}(t)$ has mass m and is subject to a force vector $\mathbf{F}(t)$ at time t then Newton's second law says

$$\begin{aligned} \text{force} &= \text{mass times acceleration} \\ \mathbf{F}(t) &= m\mathbf{r}''(t). \end{aligned}$$

Given the acceleration $\mathbf{r}''(t)$ at all times t, along with the initial position $\mathbf{r}(0)$ and the initial velocity $\mathbf{r}'(0)$, one can integrate twice to obtain the position $\mathbf{r}(t)$ at all times t.

Derivative Chain Rules. Let $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$, $g : \mathbb{R} \to \mathbb{R}$ and $f : \mathbb{R}^n \to \mathbb{R}$.

$[\mathbf{r}(g(t))]' = \mathbf{r}'(g(t))g'(t)$ [this is a vector $\mathbf{r}'(g(t))$ times a scalar $g'(t)$]

$\nabla[g(f(\mathbf{x}))] = g'(f(\mathbf{x}))\nabla f(\mathbf{x})$ [this is a scalar $g'(f(\mathbf{x}))$ times a vector $\nabla f(\mathbf{x})$]

$[f(\mathbf{r}(t))]' = \nabla f(\mathbf{r}(t)) \bullet \mathbf{r}'(t)$ [this is a vector $\nabla f(\mathbf{r}(t))$ dot product a vector $\mathbf{r}'(t)$]

Geometric meaning: A particle travels with trajectory $\mathbf{r}(t)$ through a scalar field f (say, temperature). The temperature of the particle at time t is $f(\mathbf{r}(t))$. The rate of change of temperature felt by the particle is

$$[f(\mathbf{r}(t))]' = \nabla f(\mathbf{r}(t)) \bullet \mathbf{r}'(t) = \|\nabla f(\mathbf{r}(t))\|\|\mathbf{r}'(t)\| \cos\theta.$$

This rate of change is maximized when the velocity $\mathbf{r}'(t)$ is parallel to the gradient vector $\nabla f(\mathbf{r}(t))$ (i.e., when $\theta = 0°$), minimized when the velocity is anti-parallel to the gradient ($\theta = 180°$), and is zero when the velocity is perpendicular to the gradient ($\theta = \pm 90°$).

If we write $f(x_1, \ldots, x_n)$ and $\mathbf{r}(t) = (x_1(t), \ldots, x_n(t))$ then the previous rule can also be expressed as

$$\begin{aligned} \frac{\partial f}{\partial t} &= \nabla f \bullet \mathbf{r}' \\ &= \left\langle \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle \bullet \left\langle \frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt} \right\rangle \\ &= \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}. \end{aligned}$$

The General Chain Rule (For the Curious Only!). Consider two functions $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^\ell$ with vector inputs and vector outputs, written as

$$\begin{aligned} \mathbf{f}(x_1, \ldots, x_m) &= \langle f_1(x_1, \ldots, x_m), \ldots, f_\ell(x_1, \ldots, x_m) \rangle, \\ \mathbf{g}(y_1, \ldots, y_n) &= \langle g_1(y_1, \ldots, y_n), \ldots, g_m(y_1, \ldots, y_n) \rangle. \end{aligned}$$

The "derivative" of such a function is its rectangular "Jacobian matrix":

$$D\mathbf{f} = \begin{pmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_m \\ \vdots & & \vdots \\ \partial f_\ell/\partial x_1 & \cdots & \partial f_\ell/\partial x_m \end{pmatrix} \quad \text{and} \quad D\mathbf{g} = \begin{pmatrix} \partial g_1/\partial y_1 & \cdots & \partial g_1/\partial y_n \\ \vdots & & \vdots \\ \partial g_m/\partial y_1 & \cdots & \partial g_m \partial y_n \end{pmatrix}.$$

The general chain rule says that the Jacobian matrix of the composite function $\mathbf{f} \circ \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^\ell$ equals the product of the Jacobian matrices:

$$D(\mathbf{f} \circ \mathbf{g}) = (D\mathbf{f})(D\mathbf{g}).$$

This formula includes all of the previous chain rules as special cases.

Equations of Tangent Lines and Planes. Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $F : \mathbb{R}^3 \to \mathbb{R}$.

Let $f(x_0, y_0) = c$ so the point (x_0, y_0) is on the level curve $f(x, y) = c$. The equation of the tangent line to the curve $f(x, y) = c$ at the point (x_0, y_0) is

$$\begin{aligned} \nabla f(x_0, y_0) \bullet \langle x - x_0, y - y_0 \rangle &= 0 \\ \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) &= 0. \end{aligned}$$

Let $F(x_0, y_0, z_0) = c$ so the point (x_0, y_0, z_0) is on the level surface $F(x, y, z) = c$. The equation of the tangent plane to the surface $F(x, y, z) = c$ at the point (x_0, y_0, z_0) is

$$\begin{aligned} \nabla F(x_0, y_0, z_0) \bullet \langle x - x_0, y - y_0, z - z_0 \rangle &= 0 \\ \frac{\partial F}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z - z_0) &= 0. \end{aligned}$$

The *graph* of a function $f : \mathbb{R}^2 \to \mathbb{R}$ is a special kind of surface in $\mathbb{R}^3$ defined by the equation $z = f(x, y)$. We can apply the previous formula to find the equation of the tangent plane to the graph above the point (x_0, y_0). To do this we use a **TRICK**: Define the function $F(x, y, z) = f(x, y) - z$ so that the graph $z = f(x, y)$ is the same as the surface $F(x, y, z) = 0$. In this case,

$$\nabla F = \left\langle \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right\rangle = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1 \right\rangle.$$

If $f(x_0, y_0) = z_0$ then the equation of the tangent plane at the point (x_0, y_0, z_0) is

$$\begin{aligned}
0 &= \frac{\partial F}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z - z_0) \\
0 &= \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - 1(z - z_0) \\
z - z_0 &= \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \\
z &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).
\end{aligned}$$

Linear Approximation. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ and the tangent plane to its graph over the point (x_0, y_0), which is $z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$.

If the point (x, y) is close to (x_0, y_0) then the value of $f(x, y)$ is close to the height of the tangent plane:[1]

$$f(x, y) \approx f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

We can also write this as

$$f - f_0 \approx \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

or

$$\Delta f \approx \frac{\partial f}{\partial x}(x_0, y_0)\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\Delta y.$$

If the quantities x, y are measured with uncertainties $\Delta x, \Delta y$, then the uncertainty Δf in the computed value of $f(x, y)$ has the above approximate value.

The same formula applies for any number of input variables. Consider any function $f(x_1, \ldots, x_n)$ and any point $(p_1, \ldots, p_n)$. Then for $(x_1, \ldots, x_n)$ near $(p_1, \ldots, p_n)$ we have

$$\Delta f \approx \frac{\partial f}{\partial x_1}(p_1, \ldots, p_n)\Delta x_1 + \cdots + \frac{\partial f}{\partial x_n}(p_1, \ldots, p_n)\Delta x_n.$$

Two Variable Optimization. Let $f : \mathbb{R}^2 \to \mathbb{R}$.

If f has a local max or min at the point (x_0, y_0) then we must have

$$\nabla f(x_0, y_0) = \left\langle \frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right\rangle = \langle 0, 0 \rangle.$$

[1] This is the first step of the two variable Taylor expansion.

Conversely, let (x_0, y_0) be any point satisfying $\nabla f(x_0, y_0) = \langle 0, 0\rangle$, which is called a "critical point of f". This critical point may be a local max, local min, saddle point, or something else. To tell the difference we use the "second derivative test".

Compute the "Hessian matrix" of second derivatives:

$$Hf = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}.$$

Let (x_0, y_0) be a critical point of $f(x, y)$. Then:

- If $\det(Hf)(x_0, y_0) < 0$ then (x_0, y_0) is a saddle point.
- If $\det(Hf)(x_0, y_0) > 0$ then (x_0, y_0) is a local max or min. To tell the difference, look at f_{xx}.[2] Then (x_0, y_0) is a max when $f_{xx}(x_0, y_0) < 0$ and a min when $f_{xx}(x_0, y_0) > 0$.
- If $\det(Hf)(x_0, y_0) = 0$ then the second derivative test is inconclusive. In this case we need to look at more refined information (such as the eigenvalues of the Hessian matrix, or higher derivatives), which we don't do in this course.

Change of Variables. Suppose that x and y are functions of u and v, and consider the "Jacobian determinant"

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix}$$

Then the double integrals in (x, y) and (u, v) coordinates are related as follows:

$$\iint f(x, y)\, dxdy = \iint f(x(u, v)), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv.$$

If x, y, z are functions of u, v, w then we have a similar formula

$$\iiint f(x, y, z)\, dxdydz = \iiint f(x(u, v, w), y(u, v, w), z(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| dudvdw.$$

Polar Coordinates. Let $x = r\cos\theta$ and $y = r\sin\theta$ with $0 \le r$ and $0 \le \theta < 2\pi$. Then

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix} = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r,$$

and hence[3]

$$\iint f(x, y)\, dxdy = \iint f(r\cos\theta, r\sin\theta)\, r\, drd\theta.$$

Cylindrical Coordinates. This is the same as polar, together with z. The Jacobian is

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \det \begin{pmatrix} x_r & x_\theta & x_z \\ y_r & y_\theta & y_z \\ z_r & r_\theta & z_z \end{pmatrix} = \det \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = r,$$

[2]Equivalently, you can look at f_{yy}.

[3]Since $r \ge 0$ we have $|r| = r$.

and hence we have

$$\iiint f(x,y,z)\,dxdydz = \iiint f(r\cos\theta, r\sin\theta, z)\,r\,drd\theta dz.$$

Spherical Coordinates. We will measure the point (x,y,z) by its distance ρ from the origin, its angle θ around the z-axis, measured counterclockwise from the positive x-axis, and its angle φ down from the north pole. Then we have

$$\begin{aligned} x &= \rho\sin\varphi\cos\theta, \\ y &= \rho\sin\varphi\sin\theta, \\ z &= \rho\cos\varphi. \end{aligned}$$

The Jacobian is

$$\frac{\partial(x,y,z)}{\partial(\rho,\theta,\varphi)} = \det\begin{pmatrix} x_\rho & x_\theta & x_\varphi \\ y_\rho & y_\theta & y_\varphi \\ z_\rho & z_\theta & z_\varphi \end{pmatrix} = \text{ unimportant computations } = \rho^2\sin\varphi.$$

Since $\rho \geq 0$ and $\sin\varphi \geq 0$ for $0 \leq \varphi \leq \pi$, we may take $|\rho^2\sin\varphi| = \rho^2\sin\varphi$, hence

$$\iiint f(x,y,z)\,dxdydz = \iiint f(\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi)\,\rho^2\sin\varphi\,d\rho d\theta d\varphi.$$

Integrating a Scalar Field Over a Curve. Consider a scalar field $f : \mathbb{R}^n \to \mathbb{R}$ in any number of dimensions, and consider a parametrized curve $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ in the same number of dimensions. Then we define the integral of f over the curve:

$$\int_{\text{curve}} f\,ds = \int f(\mathbf{r}(t))\|\mathbf{r}'(t)\|\,dt.$$

We can think of $ds = \|\mathbf{r}'(t)\|\,dt$ as a "tiny bit of arc length". If $n = 2$ then we can also think of this integral are the "area of a wall" with height $f(\mathbf{r}(t))$ over the point $\mathbf{r}(t) = (x,y)$ in the x,y-plane. If f is mass density then this integral is the total mass of a one dimensional wire. If $f = 1$ then this integral is just the arc length:

$$\text{Arc Length} = \int 1\|\mathbf{r}'(t)\|\,dt.$$

Integrating a Scalar Field Over a 2D Surface in 3D. Consider a scalar field $f : \mathbb{R}^3 \to \mathbb{R}$ and a parametrized surface $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$. Then we define the integral of f over the surface:

$$\iint_{\text{surface}} f\,dA = \iint f(\mathbf{r}(u,v))\|\mathbf{r}_u \times \mathbf{r}_v\|\,dudv.$$

We can think of $dA = \|\mathbf{r}_u \times \mathbf{r}_v\|\,dudv$ as a "tiny bit of surface area". To be precise, it is the area of the tiny parallelogram formed by the tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ in the u- and

v-directions. If f is a mass density then the integral is the total mass of a thin membrane. If $f = 1$ then the integral is just the surface area:

$$\text{Surface Area} = \iint_{\text{surface}} 1 \|\mathbf{r}_u \times \mathbf{r}_v\| \, dudv.$$

Integration in General (For the Curious Only!). A function $\mathbf{r} : \mathbb{R}^k \to \mathbb{R}^n$ can be thought of as a "parametrized k-dimensional shape" living in n-dimensional space.[4] We will write

$$\mathbf{r}(\mathbf{u}) = (x_1(\mathbf{u}), x_2(\mathbf{u}), \ldots, x_n(\mathbf{u})), \quad \text{where } \mathbf{u} = (u_1, \ldots, u_k).$$

Recall the definition of the Jacobian matrix:

$$D\mathbf{r} = \begin{pmatrix} \partial x_1/\partial u_1 & \cdots & \partial x_1/\partial u_n \\ \vdots & & \vdots \\ \partial x_n/\partial u_1 & \cdots & \partial x_n/\partial u_k \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{r}_{u_1} & \cdots & \mathbf{r}_{u_k} \\ | & & | \end{pmatrix},$$

whose jth column vector is the tangent vector $\mathbf{r}_{u_j} = \langle \partial x_1/\partial u_j, \ldots, \partial x_n/\partial u_j \rangle$ with respect to the jth parameter. We are also interested in the square $k \times k$ matrix $(D\mathbf{r})^T(D\mathbf{r})$, whose i, j entry is the dot product $\mathbf{r}_{u_i} \bullet \mathbf{r}_{u_j}$ of the ith and jth tangent vectors:

$$(D\mathbf{r})^T(D\mathbf{r}) = \begin{pmatrix} \mathbf{r}_{u_1} \bullet \mathbf{r}_{u_1} & \cdots & \mathbf{r}_{u_1} \bullet \mathbf{r}_{u_k} \\ \vdots & & \vdots \\ \mathbf{r}_{u_k} \bullet \mathbf{r}_{u_1} & \cdots & \mathbf{r}_{u_k} \bullet \mathbf{r}_{u_k} \end{pmatrix}.$$

Then the integral of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ over the k-dimensional region parametrized by $\mathbf{r}$ is defined as follows:

$$\int_{\substack{k\text{-dim} \\ \text{region}}} f = \int f(\mathbf{r}(u_1, \ldots, u_k)) \sqrt{\det((D\mathbf{r})^T(D\mathbf{r}))} \, du_1 du_2 \cdots du_k.$$

(Instead of a single integral sign there should be k, but that would be a mess to write.) All of the previous integral formulas are just special cases of this general formula.

Projection. The projection of a vector $\mathbf{v}$ onto the line spanned by $\mathbf{u}$ is

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \underbrace{\left(\frac{\mathbf{v} \bullet \mathbf{u}}{\|\mathbf{u}\|} \right)}_{\text{scalar}} \underbrace{\frac{\mathbf{u}}{\|\mathbf{u}\|}}_{\text{vector}}$$

Note that $\mathbf{u}/\|\mathbf{u}\|$ is always a vector of length 1. If $\mathbf{u}$ itself satisfies $\|\mathbf{u}\| = 1$ then the formula simplifies as $\text{proj}_{\mathbf{u}}(\mathbf{v}) = (\mathbf{v} \bullet \mathbf{u})\mathbf{u}$. The magnitude of the projection is the scalar

$$\mathbf{v} \bullet \frac{\mathbf{u}}{\|\mathbf{u}\|} = \text{ the component of } \mathbf{v} \text{ in the direction of } \mathbf{u},$$

[4] To avoid overlap we may need to restrict the domain. For example, the polar coordinates function $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\mathbf{r}(r, \theta) = (r\cos\theta, r\sin\theta)$ must be restricted to $r \geq 0$ and $0 \leq \theta < 2\pi$.

which is a negative number when $\mathbf{v}$ and $\mathbf{u}$ are pointing away from each other.

Integrating a Vector Field Along a Path (Line Integrals). Given a vector field $\mathbf{F} : \mathbb{R} \to \mathbb{R}^n$ and a parametrized path $\mathbf{r}(t)$, we may consider the scalar function

$$c(\mathbf{r}(t)) = \mathbf{F}(\mathbf{r}(t)) \bullet \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} = \text{component of } \mathbf{F} \text{ in the direction of } \mathbf{r}.$$

We define the integral of $\mathbf{F}$ along the path as follows:

$$\begin{aligned}
\text{Integral of } \mathbf{F} \text{ along } \mathbf{r}(t) &= \int_{\text{path}} c\, ds \\
&= \int c(\mathbf{r}(t)) \|\mathbf{r}'(t)\|\, dt \\
&= \int \mathbf{F}(\mathbf{r}(t)) \bullet \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} \|\mathbf{r}'(t)\|\, dt \\
&= \int \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt.
\end{aligned}$$

Other notations:

$$\int_{\text{path}} \mathbf{F} = \int_{\text{path}} \mathbf{F} \bullet d\mathbf{r} = \int_{\text{path}} \mathbf{F} \bullet \mathbf{T}\, ds,$$

where $d\mathbf{r} = \mathbf{r}'(t)\, dt$ is the differential of the path, $\mathbf{T} = \mathbf{r}'(t)/\|\mathbf{r}'(t)\|$ is a unit tangent vector to the path, and $ds = \|\mathbf{r}'(t)\|\, dt$ is a tiny bit of arc length.

Kinetic Energy. If a particle of mass m follows a trajectory $\mathbf{r}(t)$ in a force field $\mathbf{F}$ then Newton's second law says $\mathbf{F}(\mathbf{r}(t)) = m\mathbf{r}''(t)$. In this case, the integral of $\mathbf{F}$ along $\mathbf{r}$ is the "change in kinetic energy":

$$\int_a^b \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt = KE(b) - KE(a),$$

where the kinetic energy of the particle at time t is defined as

$$KE(t) = \frac{1}{2} m \|\mathbf{r}'(t)\|^2 = \frac{1}{2} m \mathbf{r}'(t) \bullet \mathbf{r}'(t).$$

Proof: Using the product rules for derivatives gives

$$KE'(t) = \frac{1}{2} m (\mathbf{r}'' \bullet \mathbf{r}' + \mathbf{r}' \bullet \mathbf{r}'') = m\mathbf{r}'' \bullet \mathbf{r}' = \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t),$$

and hence $\int \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt = \int KE'(t)\, dt = KE$.

Fundamental Theorem of Line Integrals. For any scalar field $f : \mathbb{R}^n \to \mathbb{R}$ we have a gradient vector field $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$. For any path $\mathbf{r}(t)$ with initial point $\mathbf{r}(0) = \mathbf{p}$ and final point $\mathbf{r}(1) = \mathbf{q}$ we have

$$\int_{\text{path}} \nabla f = f(\mathbf{p}) - f(\mathbf{q}).$$

That is, the integral $\int_{\text{path}} \nabla f$ does not depend on the "shape" of the path; only the endpoints. Proof: Recall the multivariable chain rule:

$$[f(\mathbf{r}(t))]' = \nabla f(\mathbf{r}(t)) \bullet \mathbf{r}'(t).$$

Then we have

$$\int_{\text{path}} \nabla f = \int_0^1 \nabla f(\mathbf{r}'(t)) \bullet \mathbf{r}'(t)\, dt = \int_0^1 [f(r(t))]'\, dt = f(\mathbf{r}(1)) - f(\mathbf{r}(0)).$$

Conservation of Energy. A force field $\mathbf{F}$ is called *conservative* if it is the gradient of some scalar field: $\mathbf{F} = \nabla f$ for some f. In this case we call $PE = -f$ the *potential energy* function. In this case, the total energy $KE + PE$ is conserved. Proof: As we travel along any path $\mathbf{r}(t)$ we have

$$KE'(t) = \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)$$

and

$$PE'(t) = [-f(\mathbf{r}(t))]' = -\nabla f(\mathbf{r}(t)) \bullet \mathbf{r}'(t) = -\mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t).$$

Hence

$$(KE(t) + PE(t))' = KE'(t) + PE'(t) = \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t) - \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t) = 0$$

for all times t.

Fundamental Theorem of Conservative Vector Fields. Consider a vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$. The following properties are equivalent:

- $\mathbf{F} = \nabla f$ for some scalar field f.
- For given points $\mathbf{p}$ and $\mathbf{q}$, the integral of $\mathbf{F}$ along **any** path from $\mathbf{p}$ to $\mathbf{q}$ is the same.
- The integral of $\mathbf{F}$ around any closed loop is zero.
- "Cross partials" are equal. That is, if we write

  $$\mathbf{F}(x_1, \ldots, x_n) = \langle F_1(x_1, \ldots, x_n), \ldots, F_n(x_1, \ldots, x_n)\rangle,$$

  then for all i, j we have $\partial F_i/\partial x_j = \partial F_j/\partial x_i$. In dimensions $n = 2$ and $n = 3$ we can express this condition in terms of the "curl". See the next item.

Definition of Curl. For a vector field $\mathbf{F}(x, y) = \langle P(x, y), Q(x, y)\rangle$ we define the scalar field

$$\text{curl}(\mathbf{F}) = Q_x - P_y = \frac{\partial Q}{\partial y} - \frac{\partial P}{\partial y}.$$

According to the previous item, the field $\mathbf{F}$ is conservative if and only if $\text{curl}(\mathbf{F}) = 0$. For a vector field $\mathbf{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z)\rangle$ we define the vector field

$$\text{curl}(\mathbf{F}) = \langle R_y - Q_z, P_z - R_x, Q_x - P_y\rangle.$$

According to the previous item, the field $\mathbf{F}$ is conservative if and only if $\text{curl}(\mathbf{F}) = \langle 0, 0, 0 \rangle$.

Nabla Operator. We can think of ∇ as a "vector" of "differential operators":

$$\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle.$$

This is somewhat fictional but it is useful. The gradient looks like "scalar multiplication":

$$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle.$$

And the curl of a vector field in $\mathbb{R}^3$ looks like a "cross product":

$$\begin{aligned} \text{curl}(\mathbf{F}) &= \nabla \times \mathbf{F} \\ &= \left\langle \frac{\partial f}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle \times \langle P, Q, R \rangle \\ &= \left\langle \frac{\partial}{\partial y} R - \frac{\partial}{\partial z} Q, \frac{\partial}{\partial z} P - \frac{\partial}{\partial x} R, \frac{\partial}{\partial x} Q - \frac{\partial}{\partial y} P \right\rangle \\ &= \langle R_y - Q_z, P_z - R_x, Q_x - P_y \rangle. \end{aligned}$$

Definition of Divergence. For any vector field $\mathbf{F} = \langle P, Q \rangle$ or $\mathbf{F} = \langle P, Q, R \rangle$ we define a scalar field called the divergence:

$$\text{div}(\mathbf{F}) = \nabla \bullet \mathbf{F} = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\rangle \bullet \langle P, Q \rangle = P_x + Q_y$$

or

$$\text{div}(\mathbf{F}) = \nabla \bullet \mathbf{F} = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle \bullet \langle P, Q, R \rangle = P_x + Q_y + R_z.$$

Unlike the curl, the divergence is easy to define in n-dimensions. Given a vector field $\mathbf{F} = \langle F_1, \ldots, F_n \rangle$ we define

$$\text{div}(\mathbf{F}) = \nabla \bullet \mathbf{F} = \frac{\partial F_n}{\partial x_n} + \frac{\partial F_2}{\partial x_2} + \cdots + \frac{\partial F_n}{\partial x_n}.$$

The geometric meaning of curl and divergence is expressed via the "Fundamental Theorems of Vector Calculus" (named after Green, Stokes and Gauss), as we now describe.

Green's Theorem. Let D be a region in $\mathbb{R}^2$ and let ∂D be the boundary curve, which is oriented so that D is always "to the left". Let $\mathbf{r}(t)$ be any parametrization of the boundary curve, and let $\mathbf{F} = \langle P, Q \rangle$ be any vector field. Then Green's Theorem says

$$\begin{aligned} \iint_D \text{curl}(\mathbf{F})\, dA &= \int_{\partial D} \mathbf{F} \bullet \mathbf{T}\, ds \\ \iint_D (Q_x - P_y)\, dx dy &= \int \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt. \end{aligned}$$

This theorem explains why a vector field with zero curl is conservative. Physics Proof: We can think of curl($\mathbf{F}$) as the infinitesimal amount of "counterclockwise curling" of the vector field $\mathbf{F}$. When we add up all the tiny curls, the internal curls cancel and all that is left is the component of $\mathbf{F}$ along the oriented boundary curve.

Green's Theorem is true even when ∂D consists of multiple separate loops.

Flux Integrals. The other Fundamental Theorems are expressed in terms of "flux integrals", which measure the component of a vector field $\mathbf{F}$ perpendicular to a curve $\mathbf{r}(t)$ in $\mathbb{R}^2$ or a surface $\mathbf{r}(u, v)$ in $\mathbb{R}^3$. Recall that $\mathbf{T} = \mathbf{r}'(t)/\|\mathbf{r}'(t)\|$ is a unit tangent vector to the curve $\mathbf{r}(t)$. In two dimensions we can write $\mathbf{r}(t) = \langle x(t), y(t)\rangle$ and we can use this to define a "unit normal vector"

$$\mathbf{N} = \frac{\langle y'(t), -x'(t)\rangle}{\|\mathbf{r}'(t)\|}.$$

We choose $\mathbf{N}$ so it points "to the right" of $\mathbf{r}(t)$. Then the flux of $\mathbf{F}$ across the curve $\mathbf{r}(t)$ is defined as

$$\int \mathbf{F} \bullet \mathbf{N}\, ds = \int \mathbf{F}(\mathbf{r}(t)) \bullet \frac{\langle y'(t), -x'(t)\rangle}{\|\mathbf{r}'(t)\|} \|\mathbf{r}'(t)\|\, dt = \int \mathbf{F}(\mathbf{r}(t)) \bullet \langle y'(t), -x'(t)\rangle\, dt.$$

This measures the total component of $\mathbf{F}$ that points "to the right" of $\mathbf{r}(t)$. Physically, if $\mathbf{F}$ is the velocity field of a fluid then $\int \mathbf{F} \bullet \mathbf{N}\, ds$ is the rate of flow of fluid from left to right across the curve. For a surface $\mathbf{r}(u, v)$ in three dimensions we define a "unit normal vector" by taking the cross product of the tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$:

$$\mathbf{N} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}.$$

This vector defines an orientation on the surface. Then the flux of a vector field $\mathbf{F}$ across the surface is defined as

$$\begin{aligned} \iint \mathbf{F} \bullet \mathbf{N}\, dA &= \int \mathbf{F}(\mathbf{r}(u, v)) \bullet \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|} \|\mathbf{r}_u \times \mathbf{r}_v\|\, dudv \\ &= \iint \mathbf{F}(\mathbf{r}(u, v)) \bullet (\mathbf{r}_u \times \mathbf{r}_v)\, dudv. \end{aligned}$$

Again, we can think of this as the rate of fluid flow across $\mathbf{r}(u, v)$ in the "positive direction".

Flux Form of Green's Theorem. Using the same set-up as Green's Theorem, we have

$$\iint_D \operatorname{div}(\mathbf{F})\, dA = \int_{\partial D} \mathbf{F} \bullet \mathbf{N}\, ds$$
$$\iint_D (P_x + Q_y)\, dxdy = \int_{\partial D} \mathbf{F}(\mathbf{r}(t)) \bullet \langle y'(t), -x'(t)\rangle\, dt.$$

Physics Interpretation: The total amount of expansion of a fluid inside D equals the total flow of fluid across the boundary curve. There is like "conservation of mass".

Stokes' Theorem. Let D be a two-dimensional membrane in $\mathbb{R}^3$, parametrized by $\mathbf{r}(u, v)$, and let ∂D be the boundary curve of D, parametrized by $\mathbf{r}(t)$. Then for any vector field $\mathbf{F} = \langle P, Q, R \rangle$ we have

$$\iint_D \text{curl}(\mathbf{F}) \bullet \mathbf{N}\, dA = \int_{\partial D} \mathbf{F} \bullet \mathbf{T}\, ds$$

$$\iint_D \langle R_y - Q_z, P_z - R_x, Q_x - P_y \rangle \bullet (\mathbf{r}_u \times \mathbf{r}_v)\, dudv = \int_{\partial D} \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt.$$

The proof idea is the same as Green's Theorem: We sum up the amount of (counterclockwise) curling of $\mathbf{F}$ at each point the surface. The internal curls cancel and only the boundary contribution is left. If the membrane D lives in the x, y-plane then Stokes' Theorem becomes Green's Theorem.

Divergence Theorem (Gauss' Theorem). The flux form of Green's Theorem in $\mathbb{R}^2$ becomes the Divergence Theorem in $\mathbb{R}^3$. Let D be a solid region in $\mathbb{R}^3$. Let ∂D be the boundary membrane of D with parametrization $\mathbf{r}(u, v)$ oriented to point "out of" the region D. Then for any vector field $\mathbf{F} = \langle P, Q, R \rangle$ we have

$$\iiint_D \text{div}(\mathbf{F})\, dV = \iint_{\partial D} \mathbf{F} \bullet \mathbf{N}\, ds$$

$$\iiint_D (P_x + Q_y + R_z)\, dxdydz = \iint_{\partial D} \mathbf{F}(\mathbf{r}(u, v)) \bullet (\mathbf{r}_u \times \mathbf{r}_v)\, dudv.$$

Physics Idea: The amount of expansion of a fluid inside D equals the amount of flow out of the region. This is like "conservation of mass".

What Comes Next? The language of Grad, Curl, Div and their Fundamental Theorems is the foundation of Electromagnetism. Classical mechanics also becomes simpler in this language. For example, Gauss' Law says that the divergence of the gravitational force field is equal to the density of mass that creates it. This is a more useful form of Newton's equations for gravity. The divergence of the gradient of a scalar function f is called the "Laplacian operator":

$$\nabla^2 f = \nabla \bullet (\nabla f) = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}.$$

This is in some sense the most natural generalization of the "second derivative" of a function. The Laplacian is used in the heat and wave equations in an arbitrary number of dimensions. A pressure field f evolves according to the wave equation, which is second order in time and second order in space:

$$f_{tt} = k\nabla^2 f \quad \text{for some constant } k.$$

A temperature field f evolves according to the heat equation, which is first order in time and second order in space:

$$f_t = k\nabla^2 f \quad \text{for some constant } k.$$

If the system is in equilibrium (i.e., if $f_t = 0$) then f satisfies Laplace's equation: $\nabla^2 f = 0$.